Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Engineering 38 (2012) 622 – 631

Procedia Engineering

www.elsevier.com/locate/procedia

International Conference on Modeling, Optimization and Computing (ICMOC-2012)

Nonsubsampled Contourlet Transform Based Classification of Microcalcification in Digital Mammograms

J.S.Leena Jasmine[a] , Dr.S.Baskaran[b], Dr.A.Govardhan[c]

[a] *Research scholar,ECE Department,JNTU,Hyderabad ,* [b] *Velammal Engg college, Chennai, Indiia ,*[c]*Dept. of CSE, JNTU Hyderabad 500 08, India*

Abstract

In this paper, an algorithm for the classification of microcalcification in digital mammograms using Nonsubsampled Contourlet Transform (NSCT) and Support Vector Machine (SVM) is presented. The classification of microcalcification is achieved by extracting the microcalcification features using NSCT with different scales. SVM classifier is used to classify the mammogram images based on the extracted microcalcification features. The system classifies the mammogram images as normal or abnormal, and the abnormal severity as benign or malignant. The evaluation of the system is carried on using mammography image analysis society (MIAS) database. The experimental result shows that the proposed method provides improved classification rate of over 90% for all cases.

Keywords: Mammogram;Microcalcifications; Benign; Malignant; NSCT;SVM.

1. Introduction

Breast cancer is presently one of the leading causes of death, with annual mortality of over 400,000 women worldwide. Mammography is the most common procedure to detect non – palpable cancers. Among the various radiographic indications related to the breast cancer microcalcification clusters play a vital role because they are present in 30%-50% of all cancers identified mammographically. Statistics show that 20%- 30 % of breast biopsy cases are proved cancerous. In a false negative detection, an actual tumor remains undetected that could lead to higher costs or even to the cost of a human life [1]. Essential visual clues of breast cancer include preliminary signs of masses and microcalcification clusters. The diagnosis result of tissue is classified into three categories: normal, benign and malignant. Normal represents mammogram without any cancerous cell, benign represents mammogram showing a tumor, but not produced by cancerous cells and malignant represents mammogram showing a tumor with cancerous

1877-7058 © 2012 Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
doi:10.1016/j.proeng.2012.06.077

cells. It is difficult to distinguish a benign microcalcification from one that is malignant. Consequently, many unnecessary biopsies are often undertaken due to the high positive false rate [2]

Digital mammograms are among the most difficult medical images to be read due to their low contrast and the differences in the types of tissues [3]. Thus, the task of the radiologist is tedious in the case where a significant number of mammograms require fast and accurate interpretation. To overcome this problem, a reliable automated computer-aided diagnosis system (CAD) will be very useful, providing a valuable "second opinion" to a radiologist.

Literature survey shows several techniques that have been proposed to detect and classify the presence of microcalcifications as benign or malignant using various methodologies. Chan et al. [4] developed a computerized scheme to detect the presence of microcalcifications using difference image technique in which a signal suppressed image is subtracted from a signal enhanced image to remove the structural background in the mammogram. Deepa et al [5] proposed a method using fractal background modeling, taking the difference between the original image and modeled image. Strickland et al [6] used DWT with biorthogonal splin filters to detect microcalcifications.

Karssemeijer [7], LMascio [8], J.Solka [9] developed statistical methods for the detection of microcalcifications in digital mammograms. Concerning image segmentation and specification of regions of interest (ROIs), several methods like classical image filtering and local thresholding have been proposed in [4], [10], and [11]. Furthermore, various classification methodologies have been reported for the characterization of ROI such as, rule-based systems [4], and [10], fuzzy logic systems [12], statistical methods based on Markov random fields [13] and support vector machines [14]. In addition some work reported in the literature employs neural network for cluster characterization and data mining technique for detection and classification of digital mammograms [15], [16], and [1].

Mc Leod and Parkin [17] proposed a method for detection of microcalcifications using the wavelet transform. In this method extraction of possible microcalcification is achieved by wavelet decomposition of the mammogram using daubechies wavelet to three levels. This research showed that microcalcifications are mostly prominent in the high pass sub bands of levels 2 and 3, with level 1 containing mostly noise and fine structural detail. Baguia [18] proposed new generalization of the rank nearest neighbor (RNN) rule for multivariate data for diagnosis of breast cancer. This approach suffers from a major drawback which is the unavailability of dataset comparable to Wisconsin breast cancer.

The main goal of this paper is to develop a better CAD technique for classification of microcalcification in digital mammograms using Contourlet transform and Support vector Machine. First, the features are extracted from the Contourlet coefficients which represent the unit of classification. Second, the mammogram images are classified by using Support Vector Machine (SVM). The purpose of the system is to determine the abnormal severity in the micro calcification as benign or malignant.

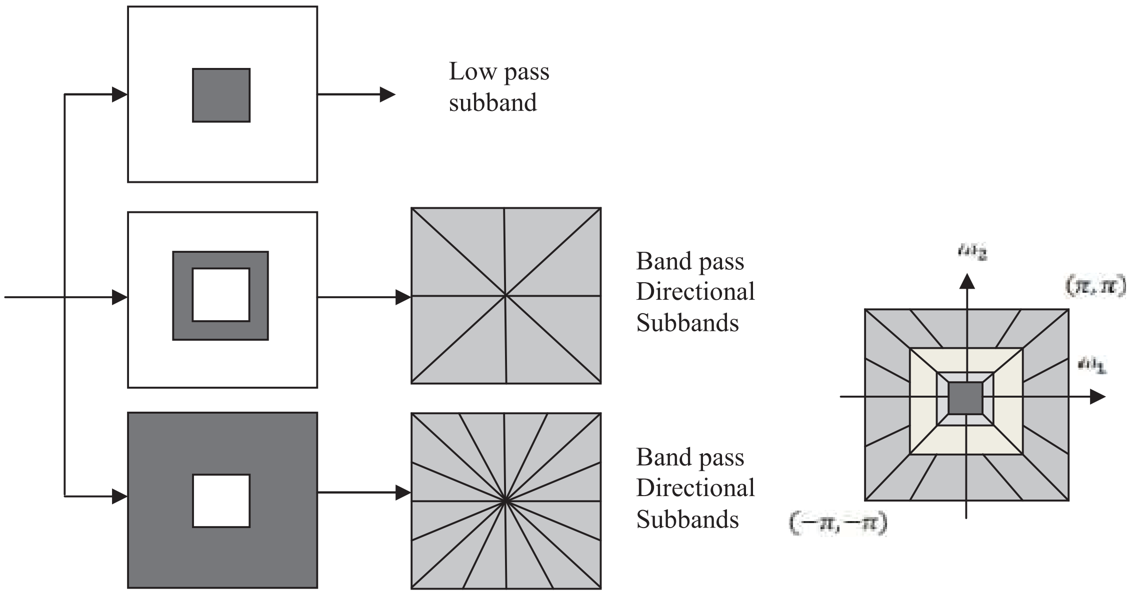
2. Methodology

The proposed system is built based on Contourlet transform of the image and by applying SVM for building the classifiers. In this section the theoretical background of both the approaches are introduced.

2.1 ContourletTransform

The Contourlet transform is an extension of the wavelet transform which uses multi scale and directional filter banks. Here images are oriented at various directions in multiple scales, with flexible aspect ratios. The Contourlet transform effectively captures smooth contours images which are the dominant feature in natural images. The main difference between Contourlet and other multi scale directional systems is that the Contourlet transform allows for different and flexible number of directions at each scale, while achieving nearly critical sampling. In addition, the Contourlet transform uses iterated filter banks, which makes it computationally efficient. The Contourlet transform [19] is a multidirectional and multi scale transform that is constructed by combining the Laplacian pyramid [20], [21] with the directional filter bank (DFB) proposed in [22]. Due to down samplers and up samplers present in both the Laplacian pyramid and the DFB, the Contourlet transform is not shift-invariant.

Fig. 1(a) displays an overview of the NSCT [19]. The structure consists in a bank of filters that splits the 2-D frequency plane in the sub bands illustrated in Fig. 1(b).This transform can thus be divided into two shift-invariant parts: 1) a Nonsubsampled pyramid structure that ensures the multi scale property and 2) a Nonsubsampled DFB structure that gives directionality. The multi scale property of the NSCT is obtained from a shift-invariant filtering structure that achieves sub band decomposition similar to that of the Laplacian pyramid. This is achieved by using two-channel non sub sampled 2-D filter banks. Fig. 2a & 2b illustrates the Nonsubsampled pyramid (NSP) decomposition with J =3 stages. The ideal pass band support of the low-pass filter at the j^{th} stage is the region [-(Π /2j), (Π /2j)]2. Accordingly, the ideal support of the equivalent high-pass filter is the complement of the low-pass, i.e., the region [(-Π /2j-1), (Π /2j-1)]2 \ [(-Π /2j), (Π /2j)]2. The filters for subsequent stages are obtained by up sampling the filters of the first stage. This gives the multi scale property without the need for additional filter design. This structure is thus different from the separable nonsubsampled wavelet transform (NSWT). In particular, one band pass image is produced at each stage resulting in J+1 redundancy. By contrast, the NSWT produces three directional images at each stage, resulting in 3J+1 redundancy.

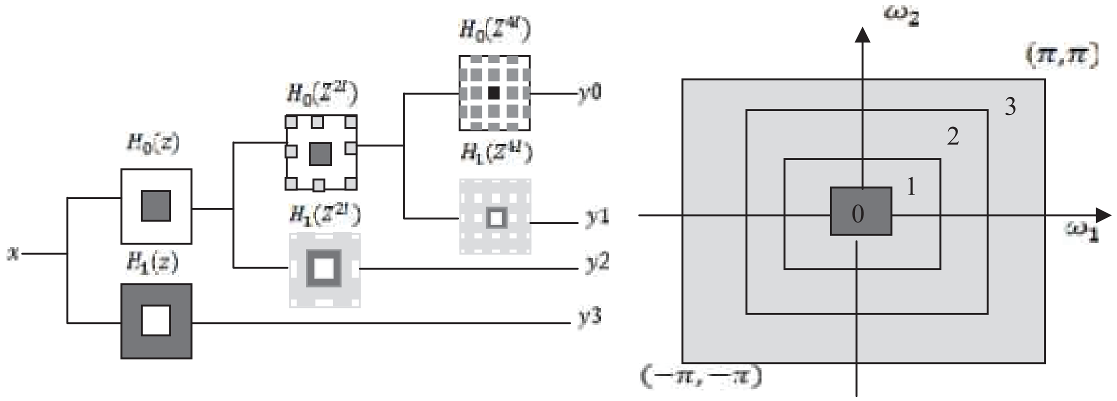


(a) (b)

Fig.1 Nonsubsampled Contourlet transform (a) Nonsubsampled filter bank structure that implements the NSCT (b) Idealized frequency partitioning.

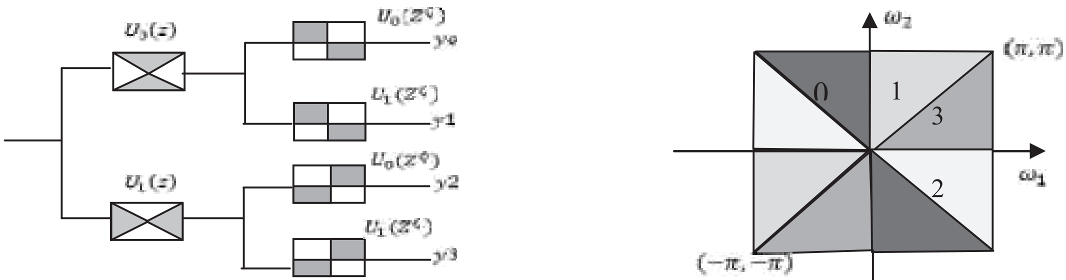
Non sub sampled Directional Filter Bank (NSDFB): The directional filter bank of Bamberger and Smith [22] is constructed by combining critically-sampled two-channel fan filter banks and re sampling operations. The result is a tree-structured filter bank that splits the 2-D frequency plane into directional wedges. A shift-invariant directional expansion is obtained with a non sub sampled DFB (NSDFB). The

NSDFB is constructed by eliminating the down samplers and up samplers in the DFB .This is done by switching off the down samplers/ up samplers in each two-channel filter bank in the DFB tree structure and up sampling the filters accordingly. This results in a tree composed of two-channel NSFBs as shown in Fig. 3a and Fig 3b illustrates the four channel decomposition. The synthesis filter bank is obtained similarly. The NSCT is flexible in that it allows any number of directions in each scale. In particular, it can satisfy the anisotropic scaling law. This property is ensured by doubling the number of directions in the NSDFB expansion at every other scale. The NSCT is constructed by combining the NSP and the NSDFB as shown in Fig. 1(a). Eight sub bands have been produced and some of the Nonsubsampled Contourlet coefficients.



(a) (b)

Fig.2.Nonsubsampled pyramid (a) Three-stage pyramid decomposition. (b) Sub bands on the 2-D frequency plane



(a) (b)

Fig.3. Four-channel Nonsubsampled directional filter bank constructed with two-channel fan filter banks. (a) Filtering structure (b) Corresponding frequency decomposition

2.2 Support Vector Machine

Support vector machines (SVMs) are a set of related supervised learning methods that analyze data and recognize patterns, used for classification and regression analysis. The standard SVM is a non-probabilistic binary linear classifier, i.e. it predicts, for each given input, which of two possible classes the input is a member of. A classification task usually involves with training and testing data which consists of some data instances. Each instance in the training set contains one "target value" (class labels) and several "attributes" (features) [24]. SVM has an extra advantage of automatic model selection in the sense that both the optimal number and locations of the basic functions are automatically obtained during training. The performance of SVM largely depends on the kernel [25].

SVM is essentially a linear learning machine. For the input training sample set

$$(x_i, y_i), i = 1 \ldots, n, x \in R^n, y \in \{-1, +1\} \quad (1)$$

the classification hyperplane equation is let to be

$$(\omega . x) + b = 0 \quad (2)$$

thus the classification margin is 2 / |ω| .To maximize the margin, that is to minimize |ω| ,the optimal hyperplane problem is transformed to quadratic programming problem as follows,

$$\begin{cases} min\phi(\omega) = \dfrac{1}{2(\omega, \omega)} \\ s.t. \; y_i((\omega . x) + b) \geq 1, \qquad i = 1,2 \ldots, l \end{cases} \quad (3)$$

After introduction of Lagrange multiplier, the dual problem is given by,

$$\begin{cases} \max Q(\alpha) = \sum_{i=1}^{n} \alpha_i - \dfrac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i y_j \alpha_i \alpha_j K(x_i, x_j) \\ s.t \; \sum_{i=1}^{n} y_i \alpha_i = 0, \qquad \alpha_i \geq 0, i = 1,2 \ldots, n \end{cases} \quad (4)$$

According to Kuhn-Tucker rules, the optimal solution must satisfy

$$\alpha_1 i \, (y_1 i \, ((w . x_1 i \,) + b) - 1 = 0, i = 1,2,..n \quad (5)$$

That is to say if the option solution is

$$\alpha^* = (\alpha_1^*, \alpha_2^*, \ldots, \alpha_i^*)^T, \qquad i = 1,2, \ldots n \quad (6)$$

Then

$$w^* = \sum_{i=1}^{n} \alpha_i^* y_i x_i$$

$$b^* = y_i - \sum_{i=1}^{n} y_i \alpha_i^* (x_i, x_j) \, , j \in \{j | \alpha_i^* > 0\} \quad (7)$$

For every training sample point x_i , there is a corresponding Lagrange multiplier. And the sample points that are corresponding to $\alpha_i = 0$ don't contribute to solve the classification hyperplane while the other points that are corresponding to $\alpha_i > 0$ do, so it is called support vectors. Hence the optimal hyperplane equation is given by,

$$\sum_{x_i \in SV} \alpha_i y_i (x_i, x_j) + b = 0 \quad (8)$$

The hard classifier is then,

$$y = sgn\left[\sum_{x_i \in SV} \alpha_i y_i (x_i, x_j) + b\right] \quad (9)$$

For nonlinear situation, SVM constructs an optimal separating hyperplane in the high dimensional space by introducing kernel function $K(x,y) = \emptyset(x).\emptyset(y)$, hence the nonlinear SVM is given by,

$$\begin{cases} min\emptyset(\omega) = \dfrac{1}{2(\omega,\omega)} \\ s.t.\, y_i\left((\omega.\emptyset(x_i)) + b\right) \geq 1,\, i = 1,2, \ldots.l \end{cases} \tag{10}$$

And its dual problem is given by,

$$\begin{cases} maxL(\alpha) = \sum_{i=1}^{l} \alpha_i - \dfrac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} y_i y_j \alpha_i \alpha_j K(x_i.x_j) \\ s.t. \sum_{i=1}^{n} y_i \alpha_i = 0, \quad 0 \leq \alpha_i \leq C, i = 1,2,\ldots,l \end{cases} \tag{11}$$

Thus the optimal hyper plane equation is determined by the solution to the optimal problem.

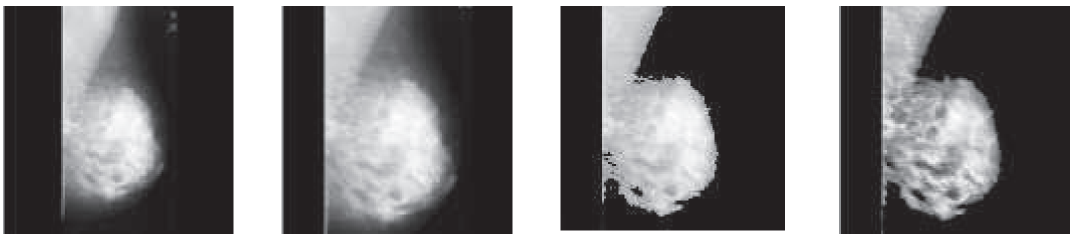
3. Proposed System

The proposed system mainly consists of two different stages which include the feature extraction and classification stage. All the stages are explained in detail in the following sub sections.

3.1 Feature Extraction stage

Feature extraction is an essential pre-processing step for pattern recognition and machine learning problems. It is often decomposed into feature construction and feature selection. In our approach, Contourlet coefficients are used as features to classify the mammogram images .The following section gives the overview of feature extraction of the digital mammogram.

In the Image pre-processing stage, the undesired distortion is suppressed and enhancement of image features and are carried out to improve the image data. The digital mammograms in the MIAS database [26], 50% of the mammograms comprised of the background with a lot of noise. To eliminate the background information and the noise, ROI image of size 800 X 800 is selected from the center of the input image. Global gray level thresholding is applied to the ROI Image. The upper threshold and lower threshold are set to 240 and 140 respectively. Then the adaptive histogram equalization is used to improve the contrast in the mammograms before feature extraction. Figure 4 shows the steps involved in the preprocessing mammograms. The Feature Extraction stage is shown in Figure 5.



(a) (b) (c) (d)

Fig.4. (a) Input image (b) ROI image (800 X 800) (c) Global thresholding (d) Enhanced image

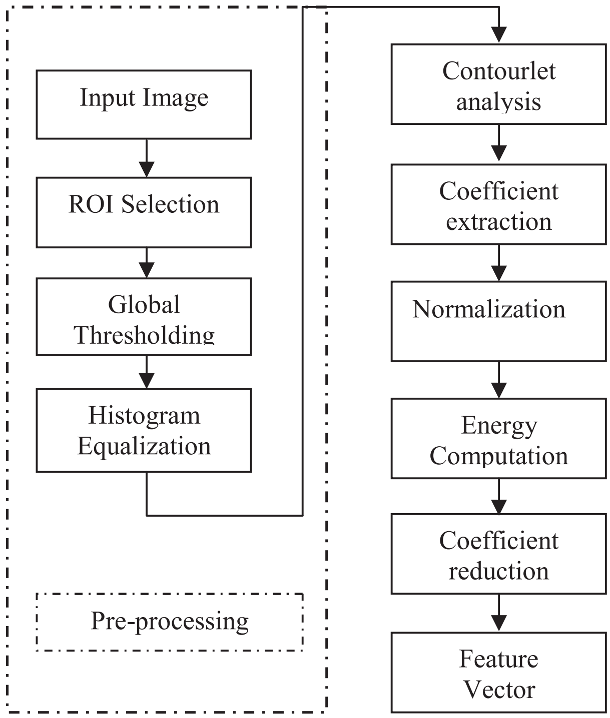


Fig.5. Feature Extraction Stage

3.1.1 Feature Vector Creation

The enhanced image is decomposed by using the NSCT at three different scales from 2 to 4. For an R level NSCT, we have 2^R directional sub bands. The Contourlet coefficients of all the sub bands are used as feature vectors individually. All the directional sub band coefficients are normalized in order to simplify the coefficient value. This is achieved by dividing each feature vector by its maximum value. Then the energy is calculated for each vector by squaring every element in the vector. The produced values are considered as features for the classification process.

As the size of ROI image is 800 x 800 and it produces high number of coefficients that are stored in a two dimensional (2D) array. To reduce the number of features by summing a predefined number of energy values together, the coefficients in 2D array is converted into 1D Array. In the proposed technique, summation of 100 and 1000 energy values per features is used. The proposed features are extracted from the training set images and stored in an array called feature vector which will be issued for training the classifier.

3.2 Classification Stage

The SVM classifier was built with two phases. In the first stage, the classifier is applied to classify mammograms into normal or abnormal categories. The mammogram is considered to be abnormal if it contains tumor (mass). If abnormal, the image enters the second stage where the abnormal mammogram is further classified into malignant or benign. The Classification Stage is shown in Figure 6.

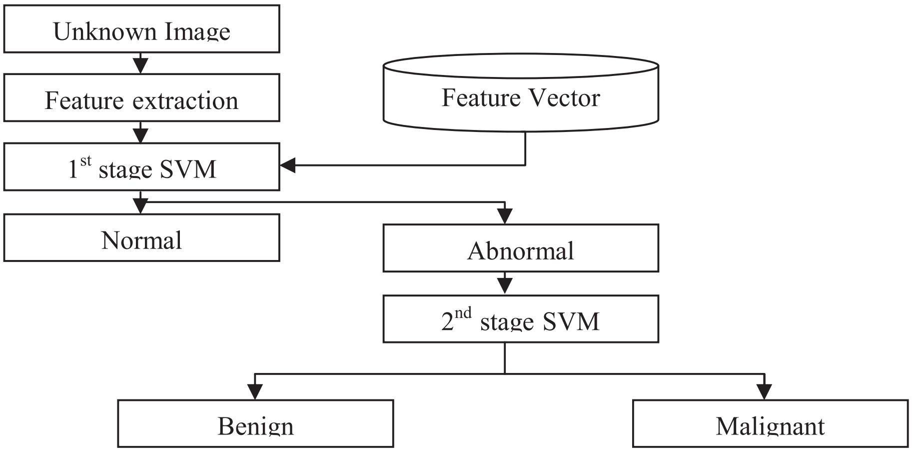


Fig.6. Classification Stage

4. Experimental Results

To assess the performance of the proposed system, many computer simulations and experiments with mammogram images were performed. The number of training and testing set is shown in Table 1. The simulations are performed by summing 100 and 1000 Contourlet coefficients per feature and trained with the 2 stage SVM classifier.

Table 1 Number of Training set and Testing set

Category	No. of Training Set	No. of testing Set
Normal	60	100
Abnormal	12	13
Micro calcification (Benign)	8	4
Microcalcification(Malignant)	4	9

The performance metric used to evaluate the accuracy of the proposed classification system is the confusion matrix. It represents information about actual and classified cases produced by a classification system. Performance of such system is commonly evaluated by demonstrating the normal and abnormal cases. Table 2 to 5 shows the confusion matrix for the proposed classification system for stage 1 and stage 2 classifier using 100 and 1000 features.

Table 2 Confusion matrixes for stage-1 classifier using 100 features

Test Outcome	NSCT Scale 2		NSCT Scale 3		NSCT Scale 4	
	Condition		Condition		Condition	
	Abnormal	Normal	Abnormal	Normal	Abnormal	Normal
Abnormal	23	10	22	9	23	9
Normal	2	90	3	91	2	91

Table 3 Confusion matrixes for stage-1 classifier using 1000 features

Test Outcome	NSCT Scale 2		NSCT Scale 3		NSCT Scale 4	
	Condition		Condition		Condition	
	Abnormal	Normal	Abnormal	Normal	Abnormal	Normal
Abnormal	24	1	23	1	20	4
Normal	1	99	2	99	5	96

Table 4 Confusion matrixes for stage-2 classifier using 100 features

Test Outcome	NSCT Scale 2		NSCT Scale 3		NSCT Scale 4	
	Condition		Condition		Condition	
	Malignant	Benign	Malignant	Benign	Malignant	Benign
Malignant	12	1	13	1	13	1
Benign	1	11	0	11	0	11

Table 5 Confusion matrixes for stage-2 classifier using 1000 features

Test Outcome	NSCT Scale 2		NSCT Scale 3		NSCT Scale 4	
	Condition		Condition		Condition	
	Malignant	Benign	Malignant	Benign	Malignant	Benign
Malignant	11	0	11	0	12	1
Benign	2	12	2	12	1	11

5. Conclusion

In this paper, an algorithm is developed for the classification of microcalcification in digital mammogram images. Nonsubsampled Contourlet transform is analyzed for features extraction, and support vector machine is used for classification process. Experimental results show that the classification rate for benign categories achieves best performance with summation of 1000 features extracted from 2 and 3-scale NSCT and summation of 100 features extracted from 3-scale NSCT for malignant categories. Our classification system produces very promising classification rate of over 90%. The evaluation of the system is carried out on MIAS dataset.

Refernecs

[1] Zaiane, O, Maria-Luiza, A & Alexandru “Application of data mining technique for medical image classification”, In proceedings of second International Workshop on Multimedia, USA.2001. pp 94-101

[2] Brigesh Verm & john Zakos, ”A computer-aided diagnosis system for digital mammograms based on fuzzy-neural and feature extraction technique”, IEEE Transactions on Information Technology in Biometrics, Volume 5, 2001 pp 46 -54.

[3] Rafayah Mousa, Qutaishat Munib & Abdallah Moussa “Breast cancer diagnosis system based on wavelet analysis and fuzzy-neural” Expert Systems with Applications, Volume 28, May 2005, pp 713–723.

[4] H.P. Chan, K.Doi, S.Galhotra, C.J.llVyborny, H.Macmahon & P.M.Jokich, “Image Feature Analysis and Computer-Aided Diagnosis in Digital Radiography”, Medi.Phy., Jul-Aug 1987,Volume 14,pp 538-548.

[5] Deepa Sankar, tessamma Thomas, “Fractal Modeling of Mammograms based on mean and Variance for the detection of Microcalcification”, IEEE Computer society, Volume 2, Dec 2007, pp 334-338.

[6] R.N.Strickland & H.I.Hahn. “Wavelet Transform for detecting microcalcification in Mammograms”, IEEE Transaction on Medical Imaging , Vol.15, April 1996, pp 218-229.

[7] Karssemeijer N, “Stochastic model for automated detection of calcifications in digital mammograms”, Image vision computing, Volume. 10, July 1992, pp 369-375.

[8] LMascio, M. Hernandez, and L. Clinton, " Automated analysis for microcalcifications in high resolution mammograms," proc SPIC – int. Soc .Opt. Eng, Jan-1993, pp. 472-479.

[9] J.Solka,W.Poston,C. priebe, G. Rogers, L. Richard , D. marchette, K.Woods and K.bowyer, " The detection of microcalcifications in mammographic images using high dimensional features," in 7 th Annual. IEEE computer based .Medical. System, June - 1993 pp. 139- 145.

[10] Davies DH, Dance DR. "Automated computer detection of clustered calcifications in digital mammogram ", Physics in Medicine and Biology. Volume.35,Aug -1990, pp.1111–1118.

[11] Qian W, Clarke LP, Kallergi M, Li H, Velthuizen R, Clark RA, et al. "Tree-structured nonlinear filter and wavelet transform for microcalcification segmentation in mammography", SPIE Biomed Image Process Biomed Visual, Volume-12,July-1993, pp509-520.

[12] Cheng H, Lui Ym ,Felimans RI, " A noval approach to microcalcification detection using fuzzy logic Technique", IEEE Transction Medical Image, Volume 17, Jun -1998, pp 442-450.

[13] Kassemeijier N, "Adaptive noise equalization and recognition of microcalcifications in mammography". Inter .J. Pattern Recog. Arti. Intel , Volume 687, 1993, pp1357-76.

[14] Bazzani A, Bevilacqua a, Bollini D, Brancaaio R, Camanini R , Lancoonelli N et al "Automated detection of clustered microcalcifications in digital mammograms using an SVM classifier", In Proceedings of the European Symposium on Artificial Neural Networks, April 2000, pp 195-200.

[15] Chan HP, Lo SCB Sahiner B, Lam Helvie MA. "Computer aided detection of mammographic microcalcifications: pattern recognition with an artificial neural network", Medical . Physics, Volume 22, Oct-1995, pg 1555- 67.

[16] Nigel RH , Nishikawa RM, Papaioannou J, Doi K, " Analysis of methods for reducing false positives in the automated detection of clustered microcalcifications in mammograms". Medical. Physics, Volume 25, Aug - 1998, pg 1502-1506.

[17] McLeod, G., & Parkin, G,"Automatic detection of clustered microcalcifications using wavelet", The third inter. Workshop on digital mammography, 199, pp 1-10.

[18] Baguia C," Breast cancer detection using rank-nearest neighbor classification rules. Pattern Recognition, Volume 36, Jan 2003, pp25-34.

[19] M.N.Do & M.Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," IEEE Trans. Image Proc., IEEE Transactions on ImageProcessing, Volume 14, Dec. 2004, pp2091-2106

[20] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," IEEE Trans. Commun., vol. 31, no. 4, pp. 532–540, April 1983.

[21] M. N. Do and M. Vetterli, "Framing pyramids," IEEE Trans. Signal Process, Volume. 51,Sep 2003, pp. 2329–2342.

[22] R. H. Bamberger and M. J. T. Smith, "A filter bank for the directional decomposition of images: Theory and design," IEEE Trans. Signal Proc., Volume 40,april 1992, pp. 882–893.

[23] Arthur L. da Cunha, Jianping Zhou, and Minh N. Do, The Nonsubsampled Contourlet Transform: Theory, Design, and Applications, IEEE Transactions On Image Processing, Volume 15,Oct 2006, pp 3089 – 3101.

[24] Y.Ireaneus Anna Rejani, S.Thamarai Selvi "Eearly detection of breast cancer using SVM classifier technique" International Journal on Computer Science and Engineering Vol.1(3), 2009, 127-130.

[25] Smola A. J., Scholkopf B., and Muller K. R., "The connection between regularization operators and support vector kernels", Neural Networks New York, vol.11, November 1998, pg 637-649.

[26] MIAS database, http://peipa.essex.ac.uk/ipa/pix/mias/